

Since 1994

Inter-University Consortium



ALMALAUREA

Connecting Universities, the Labour Market and Professionals

AlmaLaurea Working Papers - ISSN 2239-9453

ALMALAUREA WORKING PAPERS no. 4

July 2011

Integration of different data collection techniques using the propensity score

by

Furio Camillo, Valentina Conti, Silvia Ghiselli

University of Bologna

AlmaLaurea

This paper can be downloaded at:

AlmaLaurea Working Papers series

<http://www.almalaurea.it/universita/pubblicazioni/wp>

Also available at:

REsearch Papers in Economics (RePEC)

The **AlmaLaurea working paper series** is designed to make available to a wide readership selected works by AlmaLaurea staff or by outside, generally available in English or Italian. The series focuses on the study of the relationship between educational systems, society and economy, the quality of educational process, the demand and supply of education, the human capital accumulation, the structure and working of the labour markets, the assessment of educational policies.

Comments on this series are welcome and should be sent to pubblicazioni@almalaurea.it.

AlmaLaurea is a public consortium of Italian universities which, with the support of the Ministry of Education, meets the information needs of graduates, universities and the business community. AlmaLaurea has been set up in 1994 following an initiative of the Statistical Observatory of the University of Bologna. It supplies reliable and timely data on the effectiveness and efficiency of the higher education system to member universities' governing bodies, assessment units and committees responsible for teaching activities and career guidance.

AlmaLaurea:

- facilitates and improves the hiring of young graduates in the labour markets both at the national and international level;
- simplifies companies' search for personnel, reducing the gap between the demand for and supply of qualified labour (www.almalaurea.it/en/aziende/);
- makes available online more than 1.5 million curricula (in Italian and English) of graduates, including those with a pluriannual work experience (www.almalaurea.it/en/);
- ensures the optimization of human resources utilization through a steady updating of data on the careers of students holding a degree (www.almalaurea.it/en/lau/).

Each year AlmaLaurea plans two main conferences (www.almalaurea.it/en/informa/news) in which the results of the annual surveys on Graduates' Employment Conditions and Graduates' Profile are presented.

AlmaLaurea Inter-University Consortium | viale Masini 36 | 40126 Bologna (Italy)
Website: www.almalaurea.it | E-mail: pubblicazioni@almalaurea.it

The opinions expressed in the papers issued in this series do not necessarily reflect the position of AlmaLaurea

© AlmaLaurea 2011

Applications for permission to reproduce or translate all or part of this material should be made to:
AlmaLaurea Inter-University Consortium
email: pubblicazioni@almalaurea.it | fax +39 051 6088988 | phone +39 051 6088919

Integration of different data collection techniques using the propensity score

Furio Camillo[♥], Valentina Conti[♦], Silvia Ghiselli^{*}

Abstract

The present analysis has been made possible by the survey on graduates' condition that is carried out every year by the Interuniversity Consortium ALMALAUREA.

The huge number of post-reform graduates involved (about 140,000) in the survey conducted in 2008 has determined the necessity to make a radical revision of the survey methods, so that survey costs and time are reduced. This objective has been achieved through the introduction of two survey methods: CAWI (Computer-Assisted Web Interview) and CATI (Computer-Assisted Telephone Interview).

However, it is possible that the survey methods used may have influenced the answer given by graduates. In other words, since the information have been collected through different survey tools (CAWI and CATI), they may have caused distortions that are not casual.

The aim of the present analysis is to measure and to evaluate the effect of the undergone treatment on the answers given by graduates, using the propensity score approach .

In fact, many telephone and web studies have proposed the use of this approach in postsurvey impact evaluation of different interview techniques. Particularly, propensity score has been used in postsurvey weighting procedures to decrease biases.

The methodology so developed will provide a very valuable instrument for evaluating the actual reach and significance of the results presented by ALMALAUREA in its annual reports.

Keywords: survey, cawi, cati, self-selection bias, propensity score approach

1. Introduction

During the last decades, new and more powerful data collection techniques have been developed and adopted, first in the United States and, more recently, also in Italy. Indeed, compared with two decades ago when there were basically three main survey methodologies (face-to-face, telephone or e-mail interviews), today thanks to the spread of the Internet new and previously un-thought of possibilities have opened up. In particular, the CAWI technique (Computer-Assisted Web Interview) - while still representing a niche methodology - proves to be particularly significant when the target population is comprised, for the most part, of individuals characterised by high levels of educational attainment, young age on average, easy access to the Internet and good mastery of the latest technologies. A population with these characteristics, certainly a privileged one from this standpoint, can be easily accessed via the Web.

It is also true, however, that the latest Internet survey techniques, in particular the CAWI methodology, are only rarely successful in providing satisfactory response rates; for this reason, it is often necessary to resort to a combination of different survey techniques so as to guarantee that sufficient coverage and a representative sample of interviewees are achieved. Obviously, this requires researchers to solve an additional and not negligible problem, namely to make sure that the data gathered with different data collection techniques are consistent and homogeneous.

This paper sets out to verify if the responses given by interviewees submitted to different survey techniques -in this case CAWI or CATI (Computer-Assisted Telephone Interview)- can differ and, in case of biases, if the same are ascribable to the type of technique being used.

It is indeed likely that the survey methodology used can somehow affect, in a non-random manner, the responses given by graduates. For example, the presence/absence of the interviewer is a major determinant of the quality of the data collected since he/she can provide explanations or supplemental information during the interview. On the other hand, though, in some cases the interviewer's contribution may turn out to be minor or even counter-productive since he/she might influence the responses given by interviewees with his/her attitude.

[♥] Bologna University, furio.camillo@unibo.it

[♦] ALMALAUREA Inter-university Consortium, valentina.conti@almalaurea.it

^{*} ALMALAUREA Inter-university Consortium, silvia.ghiselli@almalaurea.it

2. CAWI+CATI survey for investigating the employment condition of Italian graduates

The analysis mentioned above was carried out thanks to the survey on graduates' employment conditions which the ALMALAUREA Inter-university Consortium¹ carries out on a yearly basis. This survey enables a close monitoring of the latest labour market trends by investigating the occupational outcomes of graduates from the universities participating in the Consortium during the first five years after graduation (all graduates are interviewed at one, three and five years from graduation).

More specifically, the data were derived from the last survey conducted by ALMALAUREA during 2008: they comprise, on the whole, over 287,000 interviewed graduates from 47 universities participating in the Consortium. The survey also involved all post-reform graduates of the 2007² calendar year (about 140,000) - both first and second-level – who represent the subject of this paper. The decision to involve such a large number of graduates without resorting to sampling techniques was dictated by a number of reasons; first of all, the need to meet the requirements of those responsible for the individual degree courses who need detailed documentation – at the level of each single degree course - but also of the Ministry for Education, University and Research which established that all Italian universities are bound to assess the employment outcomes of their graduates both in general and at the level of each study course. Collecting this information will serve different needs, monitoring the progress of the University Reform, but also devising and planning new study courses.

The high number of graduates involved called for the use of survey methodologies that would minimize as much as possible the costs and time of the investigation. As already mentioned above, this aim was achieved, in particular, through the adoption of a dual survey methodology, a combination of CAWI and CATI techniques. According to the survey design, the research was conducted in two different periods of the year. This was done to guarantee that the same time interval had elapsed between graduation and the interview. More specifically, those who graduated in the period between January and June 2007 were interviewed between April and June 2008, whereas those who gained their degree in the period between July and December were interviewed between September and November 2008.

¹ ALMALAUREA was set up in 1994 following an initiative of the Statistical Observatory of the University of Bologna and became a Inter-university Consortium in the year 2000. At present (as of June 2009) it brings together 53 universities accounting for about 70% of Italian graduates. ALMALAUREA was conceived as an entity operating in the service of graduates, universities and businesses. To accomplish this aim it collects and makes available reliable and updated information on the human capital trained in higher education institutions and runs a on-line database of graduates (including those with several years of experience) which is available also in English in order to facilitate the matching between labour demand and supply on the national market as well as to promote cross-border mobility. For further information on the survey methodology, see: www.almalaurea.it.

² The Italian university system underwent major changes following the Bologna Declaration (1999). The new system, which came into being in early 2000, envisages, among other things, the organisation of degree courses on two levels: a first, three-year level and a second two-year level.

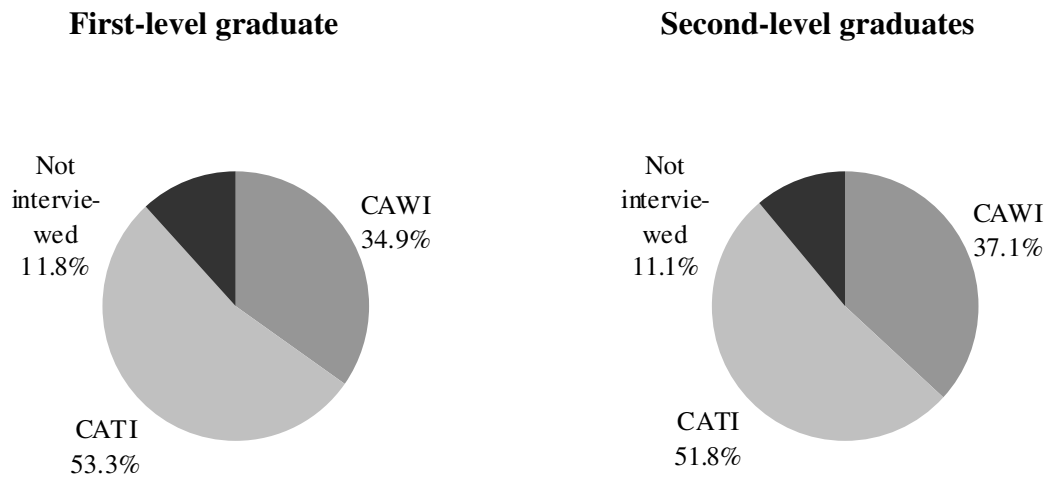


Figure 1 – Response rates of CAWI and CATI surveys by type of graduates
Data source: ALMALAUREA, year 2008.

Given the wide availability of e-mail addresses (amounting, as a whole, to 85% of post-reform graduates from the 2007 academic year), an extensive use was made of the CAWI survey (which was entirely conducted, in all its phases, within ALMALAUREA); graduates were therefore contacted by e-mail and invited to fill in a questionnaire posted on the ALMALAUREA web-site. The survey also envisaged two reminders which were sent about one week apart. The CAWI survey, which was completed within a period of about three weeks, reported unusually high response rates for this type of investigations. Response rates, as calculated in relation to the number of e-mails sent out, amounted to 41% among first-level graduates and to 43% among second-level graduates (percentages were 35% and 37% respectively, when calculated in relation to the overall surveyed population – Figure 1).

During the second phase of the survey, those who, for various reasons, had not filled in the on-line questionnaire were contacted again by phone in order to achieve satisfactory coverage and response rates for the purposes of the survey (final response rates exceeded 88% for the overall population of graduates involved in the survey, without significant differences between first and second-level graduates).

Furthermore, in order to guarantee a representative estimate of the entire population of Italian graduates (filling the gap due to the non-participation of some of the universities, as already mentioned), the findings reported in the ALMALAUREA surveys were subjected to a special statistics procedure known as “re-weighting”³.

The same survey instrument was employed in the two phases of the investigation, except for some occasional adjustments to optimise the instrument for the type of contact with the interviewee (over the Internet or by phone)⁴. The questionnaire was relatively “lean” comprising, as a whole, little

³ This is an iterative procedure whereby every graduate interviewed is given a “weight” to ensure that the variables used in the re-weighting procedure are broken down as far as possible in line with the overall population of Italian graduates. The variables considered were as follows: gender, discipline grouping, geographical area of university, place of residence of the graduate at graduation. To obtain even more accurate estimates, the cross tabulation between gender and all other variables listed above was examined (see: CISIA-CERESTA, 2001).

⁴ In particular, these adjustments consisted in changing some key words; for instance the expression “see some examples of training activities below” used in the CAWI version was changed into “I will list some examples of training activities in the following” in the CATI version. In other cases, a few minor changes were made to the

more than 30 questions, even if each graduate was able to answer a maximum of 25-27 questions, the exact number depending on the type of degree course completed (first- or second-level) and on the employment situation reported by the respondent, since graduates in work were supposed to answer a higher number of questions. The survey questionnaire proved effective for gathering the main information relating to the training and employment experiences gained by graduates after graduation; in particular, post-graduate training experiences, employment status at the time of the interview, time to entry the labour market, characteristics of the job (type of employment contract, economic sector of employment, earning, use of university-acquired skills in job, whether or not the degree was formally or substantially for the particular job), to mention but a few of the information collected.

Data relating to occupational outcomes were subsequently supplemented with the broad pool of data collected by ALMALAUREA from the questionnaires administered to students immediately before completion of their degrees. They are asked to fill in this questionnaire both for statistical purposes and for making their résumé (curriculum vitae) available to employers; questions are aimed at collecting information relating to the social-demographic characteristics of graduates (family of origin, place of residence, gender, age at graduation), pre-university studies (type of secondary school-leaving diploma, school-leaving certificate mark) and academic studies (type of degree course, degree mark, degree completion within regular time) as well as to any additional training or formal educational experiences gained during the years at university (knowledge of foreign languages and IT instruments, training periods in a company, training or study periods abroad and work experiences).

3. A preliminary descriptive analysis

Given the complexity of the investigated area, it was deemed appropriate to initially assess, by means of a simple descriptive analysis, whether or not there were structural differences between CATI and CAWI respondents.

This preliminary analysis was of fundamental importance for understanding if there is a “self-selection” of the sample of respondents to the CAWI survey in comparison to the CATI survey. In other words, if some groups of graduates are more willing than others to respond to the first stimulus as compared to the second. Should such a difference be confirmed, then it would be essential to correct for this bias before moving on to evaluate whether or not there are systematic differences related to the specific survey instrument - i.e. if there are discrepancies between the responses given in the web questionnaire and those given during the telephone interview that are solely ascribable to the type of instrument being used. In other words, if it turns out that it was in particular graduates in employment who answered the telephone interview, one would reasonably expect a higher number of “currently in work” responses in the CATI as compared to the CAWI survey. In this case, however, the bias would not be linked to the type of survey instrument, but simply to the different population of respondents.

As clearly resulted from a preliminary descriptive analysis⁵ that was carried out separately for first- and second-level graduates, those who had a good command of IT instruments, in particular web surfing skills, were generally more inclined to respond to the web survey stimulus. These were, in general, students who were fully dedicated to their university studies, regularly attended the majority of university classes and did not take on a job during their studies, unless on a temporary and part-time basis; moreover, they completed their degree with degree marks that were on average

mechanisms employed to check consistency among responses, with the specific aim of simplifying these phases in the CAWI version.

⁵ This analysis was developed through a statistic procedure called DEMOD which allows the identification of variables which, regardless of their nature, are able to characterise a given group of people. By using specific probability tests (actually chi-square tests) modalities or averages of variables (depending on the type of variable being used) are identified which are found to be significantly different from those of the overall population. Results have to be interpreted, therefore, in a probabilistic light.

higher than those observed for the overall graduate population. This particular cohort was also found to comprise younger-than-average graduates. As further proof of their high level of training they also reported, among other things, good command of English, both written and spoken, and degree marks that were generally higher than average. Both the first- and second-level graduates who had opted for the CAWI questionnaire were found to have gained their degree, more than was the case among other cohorts, in the engineering, geo-biology, sciences, agriculture or languages disciplines. Among the several other variables under investigation, it is interesting to point out that these graduates were generally more likely to be born into more advantaged families from a cultural and educational standpoint (their parents having, for the most part, a secondary school-leaving certificate or a degree title) although in some cases they had a middle school certificate; finally they were more likely to reside in the North or to have studied in a university located in the northern Italy.

By contrast, those graduates who accepted to respond to the telephone interview showed equally peculiar traits. In particular they were generally found to be of an older age, to have gained their degree with lower-than-average marks and to have been, for the most part, worker-students, i.e. individuals who were mainly dedicated to a job while pursuing university studies. For these reasons too, they usually failed to fill in the questionnaire that is submitted to them upon completion of university studies. As a result, many data relating to their experiences were not available. However, it is interesting to note that these were mainly graduates, both first- and second-level, from the healthcare professions or the law, physical education, teaching or architecture groupings. They were found to be more likely to reside in the South or to have studied in universities located in southern Italy (although there was a significant percentage of second-level degree holders who resided or had studied in the central Italy). As was to be expected, they generally exhibited poor web surfing skills.

4. Logistic regression model for assessing the variables that affect graduates' participation in the CAWI as compared to the CATI survey

After the descriptive analysis, estimating the potential bias resulting from the type of tool being used required the simultaneous assessment of all the factors that may affect the likelihood of reacting to one stimulus compared to another.

This was done through a logistic regression model which was designed to estimate the probability of self-selection between CATI and CAWI choice⁶. The variables considered related to socio-demographic data (gender, geographical area of residence, social class⁷ and parents' educational title and qualification), the academic career (discipline group, geographical area of the university, whether the graduate was *pure* or *hybrid*⁸, degree mark and age at graduation, completion of degree within the prescribed timeframe⁹, attendance of classes), skills and experiences gained during the years at university (language and IT skills, work and study periods abroad) as well as future

⁶ The model evaluates this probability using a set covariates which are selected on the basis of the preliminary descriptive analyses. To select the model researchers availed themselves of a procedure called *forward stepwise conditional process* which consists in introducing one variable at a time in the regression equation. At each stage in the process the variable having the greatest explanatory value is added to the equation. The procedure also allows for the elimination of variables that were previously introduced in the model but cease to be significant after the introduction of further explanatory factors.

⁷ Graduates were subdivided by social class following the classification proposed by A. Schizzerotto (2002). According to this method, social class is defined by the comparison between the socio-economic status of the graduate's father and mother and corresponds to the highest status of the two parents.

⁸ ALMALAUREA defines as "*pure*" those graduates who enrolled in a reformed degree course (i.e. organised on two levels); by contrast, "*hybrid*" graduates are those who have completed a reformed degree course with the contribution of educational credits gained as part of a pre-reform course. As was to be expected, this latter group of graduates is generally characterised by lower educational performance levels.

⁹ This variable considers the actual time taken by a graduate to complete university studies. This is an important aspect to analyse considering, in particular, the high number of Italian graduates (about 60% in the 2008 generation) who take more than the prescribed degree completion time to complete their academic studies.

expectations (intention to pursue further studies, willingness to move for occupational reasons, whether they would enrol again at university).

Tab. 1 – Logistic regression models - estimated by type of graduates – to evaluate the probability of participating in the CAWI survey as compared to the CATI survey

	First-level graduates			Second-level graduates		
	B	Exp(B)	Sig.	B	Exp(B)	Sig.
Gender				<i>ne</i>		
Women	<i>rm</i>					
Men	-0.102	0.903	0.000			
Geographical area of residence			0.001	<i>ne</i>		
North	<i>rm</i>					
Centre	-0.137	0.872	0.000			
South and Islands	-0.121	0.886	0.000			
Abroad	-0.011	0.989	0.912			
Parents' social class			0.000			0.000
Middle class	-0.117	0.889	0.000	-0.154	0.858	0.000
Clerical middle class	0.002	1.002	0.938	-0.058	0.944	0.126
Lower middle class	-0.093	0.911	0.000	-0.112	0.894	0.008
Working class	<i>rm</i>			<i>rm</i>		
Non responder	-0.182	0.833	0.001	-0.359	0.698	0.000
Parents' educational qualification			0.008	<i>ne</i>		
No qualification	-0.255	0.775	0.043			
Primary school certificate	-0.037	0.964	0.282			
Lower middle school cert.	-0.029	0.971	0.135			
Upper middle school cert.	<i>rm</i>					
One graduate parent only	0.022	1.022	0.326			
Both graduate parents	0.042	1.043	0.145			
Non responder	-0.221	0.802	0.004			
Study discipline group			0.000			0.000
Agriculture	-0.223	0.800	0.000	0.042	1.043	0.715
Architecture	-0.593	0.553	0.000	-0.495	0.610	0.000
Chemistry-pharmacology	-0.315	0.730	0.000	0.117	1.124	0.430
Economics-statistics	-0.344	0.709	0.000	-0.340	0.711	0.000
Engineering	-0.115	0.892	0.009	-0.067	0.935	0.337
Geo-biology	-0.190	0.827	0.000	0.034	1.035	0.665
Law	-0.546	0.579	0.000	-0.694	0.500	0.000
Languages	-0.298	0.742	0.000	-0.095	0.909	0.314
Letters	-0.443	0.642	0.000	-0.216	0.806	0.007
Medicine	-0.931	0.394	0.000	-0.322	0.725	0.008
Physical Education	-0.614	0.541	0.000	-0.475	0.622	0.001
Political-social sciences	-0.358	0.699	0.000	-0.192	0.825	0.008
Psychology	-0.346	0.707	0.000	-0.143	0.867	0.066
Sciences	<i>rm</i>			<i>rm</i>		
Teaching	-0.560	0.571	0.000	-0.390	0.677	0.001
Geographical area of university			0.000			0.004
North	<i>rm</i>			<i>rm</i>		
Centre	0.285	1.330	0.000	0.088	1.092	0.009
South and islands	0.159	1.173	0.000	-0.044	0.957	0.241
Pure/Hybrid			0.000			0.000
Pure	<i>rm</i>			<i>rm</i>		
Hybrid	0.048	1.049	0.046	0.011	1.011	0.764
Non responder	-1.918	0.147	0.000	-1.844	0.158	0.000
Degree mark	0.009	1.009	0.000	0.009	1.009	0.001
Age at graduation	-0.005	0.995	0.032	-0.011	0.989	0.004

continues

(continues) Table 1 – Logistic regression models - estimated by type of graduates – to evaluate the probability of participating in the CAWI survey as compared to the CATI survey

	First-level graduates			Second-level graduates		
	B	Exp(B)	Sig.	B	Exp(B)	Sig.
Degree completion time			0.000	<i>ne</i>		
Graduation within prescribed time, standard age	<i>rm</i>					
Graduation within prescribed time, above-standard age	-0.140	0.869	0.000			
1 year above prescribed degree completion time	-0.123	0.884	0.000			
2 years or more above prescribed degree completion time	-0.138	0.871	0.000			
Regular attendance of classes			0.000			0.019
Less than 25% of prescribed classes	-0.103	0.902	0.005	-0.119	0.888	0.094
Between 25 and 50%	-0.081	0.922	0.008	-0.132	0.876	0.050
Between 50 and 75%	-0.110	0.896	0.000	-0.113	0.893	0.006
Over 75% of prescribed classes	<i>rm</i>			<i>rm</i>		
Non responder	-0.161	0.851	0.058	-0.128	0.880	0.466
Knowledge of spoken English			0.000			0.000
None	<i>rm</i>			0.090	1.094	0.715
Limited	-0.029	0.972	0.792	0.441	1.554	0.000
Fair	-0.132	0.876	0.230	0.157	1.170	0.000
Good	-0.238	0.789	0.033	<i>rm</i>		
Very good	-0.333	0.717	0.004	-0.093	0.911	0.067
Mother tongue	-0.480	0.619	0.105	0.193	1.213	0.792
Non responder	-0.203	0.816	0.135	0.152	1.164	0.384
Knowledge of written English			0.001			0.002
None	<i>rm</i>			0.007	1.007	0.981
Limited	0.248	1.282	0.051	-0.294	0.745	0.001
Fair	0.292	1.339	0.024	-0.151	0.860	0.000
Good	0.338	1.403	0.009	<i>rm</i>		
Very good	0.414	1.512	0.002	0.050	1.051	0.283
Mother tongue	0.285	1.330	0.348	0.277	1.319	0.715
Non responder	0.133	1.142	0.395	-0.241	0.786	0.232
Web surfing skills			0.000			0.000
None	-0.556	0.574	0.000	-0.499	0.607	0.002
Limited	-0.593	0.553	0.000	-0.635	0.530	0.000
Fair	-0.450	0.637	0.000	-0.440	0.644	0.000
Good	-0.204	0.815	0.000	-0.166	0.847	0.000
Very good	<i>rm</i>			<i>rm</i>		
Non responders	-0.449	0.638	0.000	-0.267	0.766	0.012
Work experiences during academic studies			0.000			0.065
Worker-student	<i>rm</i>			<i>rm</i>		
Student-worker	0.008	1.008	0.808	0.132	1.141	0.027
No work experience	-0.070	0.933	0.064	0.173	1.189	0.007
Non responder	0.310	1.363	0.128	0.125	1.133	0.520
In work at graduation?			0.001	<i>ne</i>		
Yes	-0.068	0.934	0.000			
No	<i>rm</i>					
Non responder	-0.152	0.859	0.419			
Study-abroad experiences			0.006			0.000
Yes	0.081	1.084	0.002	0.227	1.254	0.000
No	<i>rm</i>			<i>rm</i>		
Non responder	-0.045	0.956	0.700	-0.358	0.699	0.114
Intention to pursue further studies			0.000			0.000
Yes	<i>rm</i>			0.115	1.122	0.000
No	-0.142	0.867	0.000	<i>rm</i>		
Non responder	-0.437	0.646	0.000	-0.541	0.582	0.010

continues

(continues) Table 1 – Logistic regression models - estimated by type of graduates – to evaluate the probability of participating in the CAWI survey as compared to the CATI survey

	First-level graduates			Second-level graduates		
	B	Exp(B)	Sig.	B	Exp(B)	Sig.
Willingness to move			0.000			0.000
Yes, also when it involves a change of residence	<i>rm</i>			<i>rm</i>		
Yes, also frequently	0.086	1.090	0.000	0.095	1.100	0.002
Yes, but to a limited extent	0.031	1.031	0.120	0.042	1.043	0.261
Not willing to move	-0.130	0.878	0.001	-0.204	0.816	0.029
Non responder	-0.367	0.693	0.000	-0.166	0.847	0.130
Would you enrol again at university?			0.000	<i>ne</i>		
Same degree course at the same university	<i>rm</i>					
Different degree course at the same university	0.059	1.061	0.010			
Same degree course at a different university	0.177	1.194	0.000			
Different degree course at a different university	0.096	1.100	0.001			
Would not enrol at university again	-0.146	0.864	0.022			
non responder	0.097	1.101	0.278			

Ne= variable not entered in the model

Rm= reference modality in the calculation of the *b* coefficients

Source: ALMALAUREA, year 2008.

As can be noticed, all the factors considered were structural variables and variables relating to educational and work experiences gained before graduation. The decision to include only these types of variables was based on obvious reasons, specifically the need to select “pre-treatment” factors only, i.e. those which preceded the CAWI or CATI interview.

The analysis was carried out separately for each type of graduates, meaning that two different logistic regression models were adopted for first-level and second-level degree holders. Despite the fact that a lower number of variables were included in the logistic model developed for second-level graduates, the resulting estimates were found to be very similar which confirms that self-selection of the sample of respondents depends not so much on the type of degree being completed, but rather on a different order of factors; for example, characteristics such as familiarity with the use of IT instruments and greater availability of spare time were found to be more conducive to participation to the web survey. Moreover, the findings confirmed, with a just few exceptions, the conclusions of the preliminary descriptive analysis. Any difference from the descriptive analysis that was found was due to the greater accuracy of the regression model which, as already explained, considers all factors simultaneously.

In particular, Table 1 reports the estimated coefficients for each category (columns B and Exp(B)) and the p-value for each category (column Sig.). Particularly worthy of note is the fact that among both first-level and second-level degree-holders, graduates from the sciences grouping were the most likely to respond to the CAWI survey (whereas graduates from the medicine, law, architecture, physical education and teaching groupings were less likely to opt for this mode). Indeed, it should be noted that all estimated coefficients are negative, which denotes a lower probability to respond to the web survey with respect to the reference category represented by graduates from the sciences grouping.

Moreover, as web surfing skills decrease, the probability of participating in the CAWI survey was shown to decrease as well. Indeed, compared with those who reported having very good skills, those who stated to possess no skills whatsoever were found to be 40% less likely of filling in the CAWI questionnaire.

The correct classification percentage amounted, as a whole, to 62.6% for first-level graduates and to 61.9% for second-level graduates; percentages went down to 50% in case of classification of those who participated in the web survey. The idea that self-selection of respondents depends on their pre-treatment characteristics cannot be excluded on the basis of these values which, however, are also insufficient to fully explain the self-selection mechanism. Under these circumstances it was therefore necessary to compare the sample of those who responded to the CAWI survey with the sample of participants in the CATI survey, making sure that they were placed under “equal conditions” with reference to pre-treatment characteristics. As already explained, this was done with the aim of obtaining a net measure of the bias that the use of two different survey techniques can have on the survey results.

5. Propensity score matching for evaluating differences between CAWI and CATI responses

As already anticipated, it is reasonable to expect that – all relevant conditions being the same - different approaches can give rise to different responses precisely because of the specific survey methodology being used.

Questions such as those relating to the type of employment contract, for example, may give rise to different interpretations depending on the type of approach being used. When responding to a CAWI survey, the graduate can physically scroll down and visualise on the screen all possible responses to the question and carefully evaluate what contractual form corresponds to his/her own work activity. By contrast, a telephone survey, though unable to guarantee a visual and simultaneous presentation of all response options, has one major advantage. As a matter of fact, in a telephone interview the interviewer can at any time provide the respondent with explanations to help him select the most correct contractual category.

In order to adjust for self-selection in the participation to one type of survey versus the other, respondents were classified in homogeneous groups in relation to their “propensity” to take part in the CAWI versus the CATI survey. In substance, the logistic regression model described above made it possible to assign to each graduate a probability score (a variable between 0 and 1) of participating in the CAWI survey taking into account the available pre-treatment variables: the higher the value associated to a graduate, the higher the likelihood, given the characteristics considered in the model, of participating in the web survey as compared to the telephone survey. It is therefore obvious that graduates who are assigned similar probability scores are equally similar also in terms of their propensity to respond to one type of survey versus the other. In other words, if two respondents classified within one of these groups of graduates were submitted to different approaches (i.e. different survey techniques), this can be attributed to sheer chance and not to their personal or educational experience characteristics because these aspects had already been considered in the logistic regression model. Consequently, any differences in the responses given by graduates would be attributable exclusively to the type of survey instrument being used and not to the personal characteristics of the same.

The method for evaluating the treatment-related (CAWI or CATI) differences in responses was formally developed following an approach based on impact evaluation in observational studies. Its aim consists in evaluating whether or not the type of treatment T respondents were subjected to (CAWI and CATI) exercised a significant impact on a target variable Y (the responses to the various questions contained in the administered questionnaire).

More specifically, treatment T has a causal effect on the target variable Y for the i -th unit if the result obtained in case of treatment (T_1) differs from the one obtained in the absence of a treatment (or in case of a different treatment, T_0), or if the following relation is satisfied:

$$\Delta_i \equiv Y_i(T_1) - Y_i(T_0) \neq 0$$

In reality this relation cannot be observed since each unit i is subjected either to treatment T_1 or to treatment T_0 . However, by making the right assumptions, it is possible to identify and measure a causal effect for the average individual of the population under examination.

$$E\{\Delta\} \equiv E\{Y(T_1)\} - E\{Y(T_0)\} \neq 0$$

To do so it is necessary, however, to select two independent and random samples, one being the sample subjected to treatment T_1 and the other being the control sample T_0 and to observe the target variable Y . It is true, though, that this is more difficult in the case under examination since, as already explained, the treatment itself is related to a phenomenon of self-selection of the population which, by definition, clashes with the assumption of there being a random assignment.

This problem can then be addressed by adopting the approach proposed by Rosembaum and Rubin (1983), known as *propensity score*, which is defined as the conditional probability, calculated for each unit, of receiving a treatment given a set of known covariates. It is often estimated in a logistic regression model as

$$\ln\left(\frac{e(x)}{1 - e(x)}\right) = \alpha + \beta^T f(x)$$

where $e(x) = \Pr(T_1|x)$ is the propensity score of receiving the treatment (T_1) given a set of covariates \mathbf{x} and $f(\mathbf{x})$ is some function of the covariates. For a given propensity score, covariates and study results become independent of the assigned treatment.

In summary, the logic model used in this analysis, having been adapted to the concrete circumstances being examined, is structured as follows:

X = matrix of the individual and educational experience characteristics available for the individuals, all of which relate to the pre-treatment phase and were estimated through the logistic regression model;

T = treatment variable which defines participation in one of the treatment categories (CAWI or CATI);

Y = target variable, i.e. a set of variables relating to the survey of graduates' employment conditions (for example: employment status and further training, time to entry the labour market, earnings, etc.).

Homogeneous groups of graduates were formed, therefore, on the basis of the logistic regression model. Within these groups the probability of balancing was verified, meaning the independence vis-à-vis the pre-treatment variables X ¹⁰. Verifying this probability is of fundamental importance since this is the only way one can be certain of having eliminated the self-selection bias (see Rosembaum and Rubin, 1983).

¹⁰ According to the Dehejia and Wahba (2002) binning strategy of propensity score, this feature was verified for each variable considered in the regression using a χ^2 test. For first-level graduates, the population being investigated was binned in quintiles and subsequently, having ascertained that the balancing property was not verified within the first group, the first quintile was further binned in tertiles. Within the first two tertiles the balancing properties was still not verified; for this reason this share of the population (equal to 13.3%) was not considered in the subsequent analyses. For second-level graduates, the population was first binned in quartiles and then, within the first group, in tertiles. At the end of all the necessary assessments, the first two tertiles were excluded, which corresponded to 16.7% of the overall population. Those graduates who were left out from further analyses displayed very peculiar features, so much so as to make it practically impossible to eliminate the self-selection bias linked to the type of survey and as to justify their exclusion from further in-depth analyses and investigations as outlined in the following. Moreover, these graduates were generally found not to have filled in the questionnaire that was necessary to retrieve a lot of information in relation to their pre-treatment experiences.

Within these sub-groups of population it is possible to evaluate discrepancies in the responses given by graduates and related to the different type of treatment simply by calculating the difference, for each target variable, between the observed and the expected distribution in case of absolute independence (i.e. independence between type of interview and target variable).

The resulting effects on the estimates of the response rates to the specific categories were found to be lower than 2 percentage points in terms of deviation, a very promising result since it confirms that the two survey techniques are not generating high differences in terms of responses. There were only two exceptions to this finding which enabled researchers to identify some anomalies in the way the relevant questions had been formulated in the CAWI version of the questionnaire.

The first case refers to the question on the type of job carried out: as one can infer from Table 2, the response modality “other fixed-term contract” showed a deviation equal to or higher than 2 percentage points.

This is thought to be due to the use of the word “other” which may confound the respondent, in particular when filling in the CAWI questionnaire: indeed, the word “other” may lead the graduate to consider this category to be a residual one, while in reality it is meant to refer to the actual fixed-term employment contract.

Table 2 – Observed distribution and estimate of the differences found between responses to the question relating to the type of job carried out in the CAWI survey as compared to the CATI survey, as calculated through the propensity score separately for the two types of graduates.

	First-level graduates			Second-level graduates		
	Observed distribution (column rates)	Estimate of the differences (percentage points)		Observed distribution (column rates)	Estimate of the differences (percentage points)	
		CAWI	CATI		CAWI	CATI
Permanent employment contract	30.2	1.793	-1.093	28.8	-0.543	0.462
Trainee work contract	2.9	1.000	-0.610	7.1	1.598	-1.358
Apprenticeship	6.2	0.806	-0.491	9.1	0.329	-0.279
Temporary work	2.4	0.911	-0.555	2.6	0.235	-0.199
Project work	12.1	0.274	-0.167	17.0	1.083	-0.921
Occasional collaboration	5.0	-0.753	0.459	3.5	-0.140	0.119
Provision of services	1.0	0.747	-0.455	1.4	0.842	-0.716
Socially useful/public utility work	0.1	0.092	-0.056	0.1	0.064	-0.055
Professional integration plan	0.1	0.117	-0.071	0.2	0.071	-0.060
Intermittent or on-call work	2.1	0.205	-0.125	0.5	0.063	-0.053
Job sharing	0.1	0.074	-0.045	0.0	0.027	-0.023
Auxiliary temporary work	0.5	0.580	-0.354	0.3	0.181	-0.154
Joint partnership	0.5	0.147	-0.090	0.2	0.052	-0.044
Other fixed-term contract	19.5	-3.114	1.898	20.2	-3.102	2.638
Self-employment	7.2	-1.330	0.811	5.5	-1.056	0.898
Work without any contract	9.7	-1.559	0.951	3.6	0.272	-0.231
Non responder	0.2	0.011	-0.006	0.2	0.026	-0.022

Source: ALMALAUREA, year 2008.

The second case referred to the question relating to the search for a job. Although the question is in itself very simple and unequivocal, the absence, in the CAWI version, of some consistency checks led to a certain discrepancy between responses given by graduates who had opted for the web questionnaire and responses given by those who had been interviewed by phone. For the sake of simplicity, this particular section of the questionnaire is fully reported in the following.

D1. Are you actively searching for a job? For the purposes of this survey, job searching is meant to be active which implies that at least one concrete initiative has been undertaken such as, for example, the sending of a curriculum vitae.

- [01] yes
 [02] no

[ask only if D1=01]

D2. When did you undertake the last initiative in your job searching efforts? Please remember that job searching is assumed to be active meaning that you must have undertaken at least one concrete searching action.

[01] during the last 15 days

...

[04] over 6 months ago

[05] no concrete actions have been undertaken yet. [Note for the interviewer: This is not a valid answer because job searching is meant to be active. Please go back and correct your responses] (*this modality was included only in the CATI questionnaire*)

As can be noticed, only the CATI version envisaged modality 5 in question D2: this response option could not be selected by the interviewer, its only purpose being to highlight an inconsistency between the responses given by the graduate. Indeed, how could a graduate be actively looking for a job if he has not undertaken any initiative to find one? In the CAWI version of the questionnaire, this consistency check was in fact not envisaged; in keeping with the guidelines suggested by scientific literature as well as by operational practice, researchers' goal was to make sure that filling in the questionnaire was as simple as possible and that it was formulated in a straightforward way. However, as already anticipated, the omission of this consistency check caused substantial differences in the responses given to question D1 in the two survey techniques (Table 3). However, this discrepancy did not cause remarkable problems during the data processing phase since the information contained in question D1 was used only for those graduates who had stated not to be employed (10.9% of first-level graduates and 41.3% of second-level graduates among those who participated in the CAWI survey).

Table 3 – Observed distribution and estimate of the differences found between responses to the question regarding the search for a job in the CAWI survey as compared to the CATI survey, as calculated through the propensity score separately for the two types of graduates

	First-level graduates			Second-level graduates		
	Observed distribution (column rates)	Estimate of the differences (percentage points)		Observed distribution (column rates)	Estimate of the differences (percentage points)	
		CAWI	CATI		CAWI	CATI
Looking for job	36.2	9.890	-8.020	45.5	10.705	-9.883
Not looking for job	63.8	-9.891	8.021	54.5	-10.705	9.883

Source: ALMALAUREA, year 2008.

Nevertheless, for the sake of accuracy in the analysis, researchers decided to apply a system of weights to correct and estimate the bias thus generated. According to the approach developed by Lee (2006), the adjusted weight for unit j in class c ¹¹ of the web sample becomes:

$$d_j^{w,PSA} = f_c d_j^w = \frac{\hat{N}_c^R / \hat{N}^R}{\hat{N}_c^W / \hat{N}^W} d_j^w$$

where

¹¹ This refers to each group within which the balancing property was verified.

d_j^w = base weight, necessary in order to get representative estimates of the Italian population

\hat{N}_c^R = number of units in class c in the reference survey (in this case, the CATI survey)

\hat{N}^R = total number of units in the reference survey (in this case, the CATI survey)

\hat{N}_c^W = number of units in class c in the web survey

\hat{N}^W = total number of units in the web survey

It should be noted that each graduate of each group c was assigned the same weight f_c .

Table. 4 – Employment status by type of graduate: comparison between weighted values and weighted and adjusted values using the weight system proposed by Lee

	Employment status	
	Weighted values	Weighted and adjusted values
First-level graduates		
In work	29.2	31.5
In work and enrolled on the second-level degree course	16.1	16.0
Enrolled on the second-level degree course (and not in work)	44.6	42.0
Not looking for work	2.8	2.9
Looking for work	7.3	7.6
Second-level graduates		
In work	61.7	62.1
Not looking for work	17.7	17.4
Looking for work	20.6	20.5

Source: ALMALAUREA, year 2008.

Table 4 reports the results obtained for graduates' employment status by using the base weight d_j^w (first column), which is necessary in order to get representative estimates of the Italian population, and the adjusted weight $d_j^{w,PSA}$ (second column). As can be noticed, differences are decidedly small, in the order of just a couple of percentage points. This confirms that, despite the problems in retrieving the information regarding the search for a job, the actual bias is, in fact, not very relevant, according to the usual sampling error rate results in large sample surveys.

6. Conclusions

After a brief descriptive analysis, this paper shows the application of the propensity score matching method to evaluate whether there are any differences in the responses, to the same questions, given to CAWI as compared to CATI surveys. In this specific case, the analysis was focused on about 140,000 Italian graduates from the 2007 calendar year who had participated in the 2008 survey that the ALMALAUREA Consortium conducts on a yearly basis to evaluate graduates' employment conditions.

Two separate analyses were carried out for each category of degree holders (first- or second-level graduates) and no significant differences emerged between the two cohorts. This means that self-selection of the respondents' sample is linked not so much to the type of degree title but, rather, to a series of other factors concerning, for the most part, familiarity with the use of IT instruments and availability of spare time. These factors were found to significantly increase graduates' chances of participating in the web survey.

Logistic regression models were used to identify those variables which have the strongest impact on graduates' propensity to respond to the web survey as compared to the telephone survey. It emerged that graduates from the sciences groupings are more likely to respond to the web survey as opposed to those from the medicine, law, architecture, physical education and teaching discipline groups. Moreover, to mention just the most interesting findings, the chances of participating in the web

interview were found to become higher with increasing IT skills meaning, in particular, the ability to navigate the web with ease and competence.

Thanks to the logistic regression models it was also possible to attribute to each graduate a probability of participation in the CAWI survey on the basis of the available pre-treatment variables. This led, in a later phase, to a break-down of graduates in homogeneous groups based on their probability score. This means that if, within a given group of graduates, one individual had participated in the web survey and another in the telephone survey, this could be ascribable only to sheer chance since the two individuals were found by the regression model to be comparable in terms of probability. For this reason it was possible to evaluate differences between the responses given in the CAWI survey and those given in the CATI interview within each group of graduates.

Discrepancies in the responses provided to the questionnaires were found to be decidedly low, in the order of just a few percentage points, with two notable exceptions which made it possible to identify two anomalies in the formulation of the CAWI questionnaire. One such discrepancy concerned the question regarding the type of job carried out where the use of the word “other” in one of the response modalities in the CAWI version had led to an underestimate of this option by respondents. In the second case, by contrast, the omission of a consistency check between some responses given by interviewees had actually led to an increase in the share of graduates who stated to be searching for a job. This is likely to be ascribable in particular to the absence of the interviewer, given his/her very important role in providing clarifications and explanations, as necessary, during the interview. As for future surveys, the ALMALAUREA Consortium it advised to correct these anomalies in the CAWI version of its questionnaire. Luckily, these problems did not give rise to significant errors during the data analysis phase: discrepancies were namely in the order of just about 2 percentage points. To account for this discrepancy, a system of weights was applied which made it possible to correct the bias ascribable to the survey instrument.

The methodology so developed, as outlined in this paper, will provide a very valuable instrument for evaluating the actual reach and significance of the results presented by ALMALAUREA in its annual reports. Its application potential will further increase in the future since the Consortium is expected to make greater use of a combination of the two survey instruments – namely CAWI and CATI – in order to be able to collect a wide range of data at relatively low costs.

In the future this evaluation will be further developed and enriched through the application of techniques to create comparability conditions among units subjected to different survey methodologies which make use, in multivariate terms, of all the information available in the pre-treatment set of variables. Such techniques (Camillo, D’Attoma 2009) also enable researchers to generate a real semi-automatic impact evaluation system since they exploit, using a data mining and entirely data-driven approach, the geometric properties of the multivariate spaces generated by the possible interactions (if any) between pre-treatment variables. Preliminary analyses¹² have led to promising results confirming the assumptions contained in this paper.

¹² This refers to an in-depth analysis presented by the authors during ITACOSM09, First Italian Conference on Survey methodology, held in Siena on June 10-12, 2009.

References

Camillo F., Girotti C. (2007): L'impatto dell'integrazione di tecniche multiple di rilevazione nell'indagine sulla condizione occupazionale AlmaLaurea: una misura di propensity score in spazi condizionati multivariati. In: IX Profilo dei laureati italiani. La riforma allo specchio, Consorzio Interuniversitario AlmaLaurea (edited by), Il Mulino, Bologna, 289-309.

Camillo F., D'Attoma I. (2009): A New Data Mining Approach to Estimate Causal Effects of Policy Interventions, Expert Systems with Applications doi:10.1016/j.eswa.2009.05.072

Cammelli A. (2009): XI rapporto sulla condizione occupazionale dei laureati. Occupazione e occupabilità dei laureati a dieci anni dalla Dichiarazione di Bologna, available on: www.almalaurea.it/universita/occupazione/occupazione07.

CISIA-CERESTA (2001): Manuale di SPAD. Version 4.5, Paris.

Dehejia R. and Wahba S. (2002): Propensity score-matching methods for nonexperimental causal studies, The Review of Economics and Statistics, 84(1), 151-161.

Lee S. (2006): Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys. Journal of Official Statistics, Vol. 22, No 2, 329-349.

Rosenbaum, P.R. and Rubin, D.B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika, 70, 41-55.

Rosenbaum, P.R. and Rubin, D.B. (1984): Reducing Bias in Observational Studies Using Subclassification on the Propensity Score, Journal of the American Statistical Association, 79, 516-524.

Schizzerotto A., edited by, (2002): Vite ineguali. Disuguaglianze e corsi di vita nell'Italia contemporanea, Bologna, Il Mulino.